REDACT: PII Redaction and Privacy Protection with Ollama Integration is the proposed title of this paper

Question: How to achieve privacy?

Answer: Privacy in the proposed system is achieved through a combination of secure data handling, advanced redaction techniques, and architectural decisions that prioritize confidentiality. All data is processed locally within a secure environment, ensuring no sensitive information is transmitted to external servers or third-party APIs. Textual redaction is performed using PyMuPDF, which applies irreversible black box redaction to structured PDFs, while OpenCV handles visual redaction by detecting and obscuring faces and other visual identifiers in scanned documents and images. PII detection is carried out using open-source large language models hosted locally via Ollama, eliminating the need for cloud-based AI services and further enhancing privacy. Together, these techniques ensure compliance with major data protection regulations such as GDPR, HIPAA, and CCPA, making the system secure, private, and legally compliant.

Question: How to improve the protection process?

Answer: The protection process can be improved by incorporating several enhancements focused on accuracy, flexibility, and contextual awareness. One key improvement is the implementation of priority-level redaction, which allows users to assign varying levels of sensitivity to different types of PII—such as treating names as low priority and Aadhar numbers as high priority—offering finer control over what and how information is redacted. Additionally, enhancing the Faker module to generate more region-specific and realistic synthetic data would strengthen anonymization, especially for documents subject to localized regulations. The integration of advanced natural language understanding models can also improve the contextual identification of PII, reducing false positives and negatives. Furthermore, supporting handwritten text recognition and expanding multilingual capabilities would make the system more versatile. Finally, incorporating encryption mechanisms for redacted documents and logs ensures that even intermediate data remains secure, further strengthening the overall protection process.

Question: Clarity needed for optical character recognition.

Answer: Optical Character Recognition (OCR) is a crucial component of the system that enables the extraction of text from non-editable sources such as scanned PDFs and images. In this solution, Tesseract OCR is used due to its high accuracy and open-source nature. The OCR engine analyses the visual structure of the document, detects characters and words, and

converts them into machine-readable text. This process ensures that even handwritten or printed text in scanned documents can be processed further for PII detection. By accurately converting unstructured visual data into structured text, OCR acts as a bridge between raw image-based content and intelligent data processing, making it possible to identify and redact sensitive information effectively across various document formats.

Question: Survey of literature is not systematic

Answer: Revised the literature review section to ensure a more systematic and structured presentation. Kindly refer to the revised section in the paper for the changes.

Question: How to achieve reliability?

Answer: The application achieves reliability by integrating robust technologies such as Tesseract OCR for accurate text extraction and advanced open-source large language models from Ollama to ensure consistent identification of sensitive data across various document types and languages. OpenCV-powered image processing enhances the detection and redaction of visual PII, including faces. Reliability is further reinforced through extensive validation using diverse sample documents, including edge cases, to minimize false positives and negatives. Built-in logging and monitoring systems continuously track redaction accuracy and system health. The application allows users to review and modify redacted content before finalization, providing an additional layer of verification that reduces the risk of error. Continuous updates to redaction logic and retraining of language models based on real-world feedback help maintain consistent performance and trustworthiness over time.

Question: How to improve performance?

Answer: The application improves performance by leveraging efficient libraries such as PyMuPDF for rapid PDF parsing and using Tesseract OCR with fine-tuned settings to accelerate text extraction. GPU acceleration is utilized for both OCR and LLM inference, significantly reducing latency when processing large batches of documents. Locally hosted open-source large language models are optimized through quantization and model distillation, ensuring high-speed PII detection without compromising accuracy. The system employs parallel processing and asynchronous workflows, enabling simultaneous handling of multiple documents. Intermediate results are cached, and redundant computations are eliminated to streamline operations. Documents are processed in batches, and the pipeline is modularized, allowing each component to scale independently based on workload. Continuous profiling and

tuning ensure that the application maintains high scalability and responsiveness, even under heavy usage.

Question: How to achieve security?

Answer:The application achieves security by ensuring that sensitive data is processed entirely within a secure, controlled environment without being exposed to external systems. It leverages Ollama's open-source large language models, which operate locally and do not store, transmit, or manipulate user data beyond the immediate session. This eliminates the risk of data leaks associated with cloud-based AI services. The application ensures that no sensitive information leaves the system by running all processing tasks—including OCR, PII detection, and redaction—on the user's local infrastructure. It enforces strict access control through authentication and role-based permissions, ensuring that only authorized users can interact with or view the redacted content. Redaction is performed using irreversible techniques like black-box overlays and image masking to make PII unrecoverable. Temporary files are purged automatically after processing, and detailed logs are maintained to monitor activity and identify any anomalies. Regular vulnerability assessments and compliance with privacy regulations like GDPR and HIPAA reinforce the application’s commitment to secure, private, and responsible data handling.

Question: How to achieve data protection?

Answer: The application achieves data protection through a combination of secure processing, encryption, controlled access, and compliance-driven design. All sensitive data is handled within a secure, isolated environment, ensuring it never leaves the local system or gets exposed to third-party services. Access to the system is restricted using authentication mechanisms and role-based permissions, ensuring only trusted users can view or process sensitive content. Irreversible redaction methods are applied to remove PII permanently from both text and images, preventing any possibility of recovery. The entire workflow aligns with global data protection standards such as GDPR and HIPAA, offering strong assurance of confidentiality, integrity, and compliance throughout the data lifecycle. In the future, the system will incorporate the Faker module to replace redacted PII with context-aware synthetic data, further enhancing privacy while maintaining the readability and structure of documents.

Comment: It is suggested to organize the Conclusion and future scope section much better.

Response: Reorganised the conclusion and future scope section to improve clarity. Kindly refer to the updated section.

Reviewer 2 Comments:

Comment:
1. Results section is not strong enough. Provide more conceptual details about obtained in.
2. It Should include numerical values for the results, comparing numerically with different methods.

Response: Revised the results section with numerical values for better understanding. Kindly refer to the updated section.

Comment:
3. There is no logical derivation of research workflow.

Response: The workflow has been given in detail under the title ‘Proposed Solution’ in the paper. Kindly refer to it.

Comment:
4. Authors are asked to provide a comparison chart or table with the existing algorithm in the results and discussion section.
5. Need a real graph (accuracy and loss) and explanation of it in the result section. Provide a real time graph that is given by simulation software. Here no real time graph included.

Response: Added a detailed comparison table that contrasts the performance of our proposed method with existing approaches, including manual redaction and a traditional machine learning model. Kindly refer to it in the paper.